



Education Policy Analysis
Archives/Archivos Analíticos de Políticas
Educativas

ISSN: 1068-2341

epaa@alperin.ca

Arizona State University
Estados Unidos

Neely, Stephen R.; Diebold, Jeffrey
Public Expenditures and the Production of Education
Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 24,
2016, pp. 1-18
Arizona State University
Arizona, Estados Unidos

Available in: <http://www.redalyc.org/articulo.oa?id=275043450081>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



epaa | aaape

Arizona State University

Volume 24 Number 88

August 15, 2016

ISSN 1068-2341

Public Expenditures and the Production of Education

Stephen R. Neely

University of South Florida



Jeffrey Diebold

North Carolina State University
United States

Citation: Neely, S.R. & Diebold, J. (2016). Public expenditures and the production of education. *Education Policy Analysis Archives*, 24(88). <http://dx.doi.org/10.14507/epaa.24.2441>

Abstract: The relationship between public education expenditures and student outcomes remains an important concern for policy analysts, educational administrators, and the public at large. While previous studies have failed to identify a consistent relationship between public investments in education and positive student outcomes, most analyses have not accounted for the different educational goals associated with various instructional expenditure categories. This study builds on prior research by using Pennsylvania's public school districts to test proposed improvements in model specification for the traditional education production function. Using longitudinal, fixed-effects models, a detailed disaggregation of instructional expenditures is undertaken in order to account for the likelihood that different instructional subcategories (i.e. regular programming, special education, and vocational instruction) influence student outcomes in varying ways. The results suggest that the impact of expenditures may be understated in previous studies based on a failure to account for these distinctions, particularly in the case of Mathematics education.

Keywords: education production function; school finance; education policy; education finance

El gasto público y la función de producción de la educación

Resumen: La relación entre el gasto en educación pública y los resultados de los estudiantes sigue siendo una preocupación importante para los analistas de políticas, administradores de la educación, y el público en general. Mientras que los estudios anteriores no han logrado identificar una relación consistente entre la inversión pública en educación y resultados positivos de los estudiantes, la mayoría de los análisis no han contabilizado los diferentes objetivos educativos relacionados con las distintas categorías de gastos educativos. Este estudio se basa en investigaciones anteriores en distritos escolares públicos de Pensilvania para comprobar las mejoras propuestas en la especificación del modelo tradicional de la función de producción de la educación. Utilizamos un análisis de modelos de efectos fijos longitudinales, a través de un desglose detallado de los gastos de instrucción, con el fin de tener en cuenta la probabilidad de que las diferentes subcategorías de instrucción (programación regular, educación especial, e instrucción profesional) influyen en los resultados de los estudiantes en diferentes maneras. Los resultados sugieren que el impacto de los gastos podría haber sido subestimada en estudios previos basados ya que no tomaron en cuenta estas diferencias, sobre todo en el caso de la educación matemática.

Palabras clave: función de producción de educación; financiamiento escolar; política educativa; financiación de la educación

Gastos públicos e a função de produção da educação

Resumo: A relação entre gastos na educação pública e os resultados dos alunos continua a ser uma grande preocupação para analistas políticos, administradores educacionais e o público em geral. Embora estudos anteriores não tenham conseguido identificar uma relação consistente entre o investimento público na educação e resultados positivos para os alunos, a maioria das análises não consideraram os diferentes objetivos educacionais relacionadas com as várias categorias de despesas educacionais. Este estudo baseia-se em pesquisas anteriores em distritos escolares públicos na Pensilvânia para verificar as melhorias propostas na especificação do modelo tradicional da função de produção da educação. Usamos uma análise do modelo de efeitos fixos longitudinais, através de um detalhamento dos gastos de instrução, a fim de tomar em conta a probabilidade de que diferentes subcategorias de instrução (programação regular, educação especial e educação profissional) influenciam os resultados de estudantes de diferentes maneiras. Os resultados sugerem que o impacto dos gastos poderiam ter sido subestimados em estudos anteriores já que não tiveram em conta estas diferenças, especialmente no caso da educação matemática.

Palavras-chave: função de produção de educação; financiamento da escola; política de educação; financiamento da educação

Introduction

Among the chief ambitions of public school finance is the organization of schools and the allocation of resources in those ways which maximize positive student outcomes. Whether gauging the potential impact of new expenditures or reconsidering the current distribution of resources, public policymakers and local administrators alike are interested in how scarce, limited inputs can be most efficiently applied to the attainment of educational goals. To this end, analysts have often adopted the language and methodologies of manufacturing and production in order to measure the “value” of resource inputs in the educational process. Developing econometric models called *education production functions*, researchers have applied the logic of the factory to a variety of education policy concerns, such as the impact of class-size reductions on student outcomes, the importance of

quality teachers, and the predicted influence of increased educational spending on student test-scores.

While this approach offers the potential for improving our understanding of educational productivity (Monk, 1989), the cumulative research up to this point has been inconclusive at best, and at times contradictory. Rice and Schwartz (2008) note that studies examining the relationship between public expenditures and student performance have been “frustratingly inconsistent in their findings” (p. 136). For example, while some comprehensive analyses have argued that there is a clear relationship between education expenditures and productivity (i.e. Greenwald, Hedges, & Laine, 1996; Krueger, 2002), others have concluded that increases in education funding are unlikely to produce any measurable improvements in student outcomes (i.e. Hanushek, 1997, 2003). The latter conclusion in particular has become nearly axiomatic in many policy circles (see Baker, 2012, for discussion).

In an effort to better understand the relationship between public expenditures and student outcomes, subsequent analyses have pursued enhancements in model specification through various approaches such as the disaggregation of expenditure variables and the use of “value-added” outcome measures. Following this logic, several researchers have specifically attempted to disaggregate financial inputs into major expenditures categories (such as instructional and support service expenditures), though even these analyses have struggled to achieve a common consensus (i.e., Dee, 2005; O’Connell Smith, 2004; Wenglinsky, 1997).

As economic challenges continue to tighten state budgets, understanding the relationship between public expenditures and educational production becomes increasingly important at both the macro and micro policy levels. For example, future investments in education will most likely be influenced by the extent to which policymakers (and their constituents) feel that current levels of investment are impacting student outcomes. To that end, this article builds on previous production function studies by more closely examining the specific types and uses of instructional expenditures and how they influence student outcomes. Unlike many previous studies, this analysis accounts for the differential impacts that may be anticipated from various expenditure categories, such as the impact of regular program expenditures on standardized test scores, versus that of special needs or vocational education expenditures. By accounting for these more nuanced differences, this research seeks to provide policymakers with a more complete understanding of the relationship between public expenditures and educational outcomes.

While the effect sizes found in this analysis are generally small, they do suggest that the relationship between per pupil expenditures and educational productivity may have been understated in some previous studies based on a failure to account for various types of instructional expenditures. This conclusion is particularly salient in the case of Mathematics education, where the disaggregation of instructional expenditures into specific subcategories results in a larger measured impact on student test scores for *regular program expenditures* than for the aggregate *instructional expenditures* category. In contrast, when these expenditures are not properly disaggregated, the impact of instructional expenditures appears smaller due to the inclusion of expenditures primarily associated with different educational goals, such as special needs and vocational education. While caution is suggested with regard to over-interpreting the parameter estimates of district-level analyses, the findings do demonstrate a general relationship which warrants further attention from both researchers and policymakers.

The Production Function Model

Production function research applies the logic of manufacturing firms to the production of educational outcomes in an effort to better understand the relationships between resource inputs/allocations and student outcomes. Monk (1989) notes that

... a production function ... describes the maximum level of outcome possible from alternative combinations of inputs. It summarizes technical relationships between and among inputs and outcomes. The production function tells what is currently possible. It provides a standard against which practice can be evaluated on productivity grounds. (p. 31)

Previously, researchers have pointed out the challenges of applying a production metaphor to educational concerns (see Summers & Wolfe, 1979), but since the publication of the Coleman Report (1966), production function analysis has become a mainstay in school finance and education policy research. The essential elements of a production function model vary based on the observed units of analysis. For school district-level aggregation (which this article employs) the general form of the production function model is essentially as follows:

$$Y_{it} = f(S_{it}, P_{it}, O_{it}, E_{it}) \quad \text{Eq. (1)}$$

where Y_{it} represents a measure of student outcomes for i district at time t ; S_{it} represents measures of school resource inputs for i district at time t ; P_{it} represents relevant measures of “peer-group” or student-body characteristics (such as race and socio-economic status) for i district at time t ; O_{it} represents organizational characteristics (such as school size) for district i at time t ; and E_{it} represents environmental characteristics (such as urban or rural settings) for district i at time t .

By controlling for factors that are known to influence student outcomes, such as peer-group and environmental characteristics, production functions allow researchers to capture the impact of those variables which lie within the influence of policymakers, such as the quantity and allocation of school resources. In essence, the production function allows the effect of school resources to be isolated, highlighting the anticipated marginal impact of a one unit increase in a given resource. Interpretation of production function models requires caution, as the analysis cannot account for unobserved factors such as student effort (Baird, 2011). However, production function analysis does allow for a general understanding of the direction and magnitude of relationships between resource inputs and those student outcomes which are most desirable from a policy standpoint.

Previous Studies

The initial focus on resource inputs in education was derived from early 20th century, closed-system organizational theories, which emphasized the role of internal processes in the production of organizational outcomes (Marion & Flanigan, 2001). In the field of education, these long held assumptions were challenged by the publication of the Coleman Report in 1966, which concluded instead that individual and environmental factors, such as socioeconomic status and student background, were the primary determinants of educational outcomes. After conducting a large scale production function analysis on behalf of the U.S. Office of Education, Coleman et al. (1966) concluded that “... schools bring very little influence to bear on a child’s achievement that is

independent of his background and general social context” (Coleman et al., p. 325). These findings led to a significant shift in thinking on the part of many policymakers, suggesting that further public investments in education may be in vain given the prevailing influence of social and environmental factors. It should be noted that the Coleman Report was met with a healthy degree of skepticism on the part of many scholars. For example, in one influential critique, Bowles and Levin (1968) argued that the analysis underlying Coleman et al.’s (1966) conclusions was lacking in a number of areas, including (1) a disproportionate response rate that heavily weighted suburban schools, (2) inadequate analysis and treatment of non-responses, (3) poor operational measurement of school resources, and (4) a limited, cross-sectional research design.

However, despite these and other criticisms, the Coleman Report’s findings sparked a robust debate, which led many researchers to reexamine his claims from a variety of sampling and methodological approaches. Within 30 years of the Coleman Report’s initial publication, Hanushek (1997) was able to conduct an analysis of 377 published production function estimates, and he found strong support for the Coleman Report’s initial conclusions. Of the 377 estimates analyzed by Hanushek, 163 examined per pupil expenditures as an input variable, and they were only found to be a statistically significant predictor of positive student outcomes 27% of the time. While Hanushek’s “vote-counting” method has been criticized by several researchers (i.e. Greenwald, Hedges, & Laine, 1996; Krueger 2002), his analysis became widely cited in many policy circles as evidence that public expenditures do not significantly contribute to improved educational outcomes (see Baker, 2012, for discussion).

Yet while this “non-relationship” between financial inputs and student outcomes has become a popular theme in many policy debates, a number of subsequent studies have sought to challenge this assertion through improved model specification. One attempted avenue of improvement has been the use of disaggregated expenditure categories, which stems from the assumption that dollars spent on different functions should have distinct impacts on productivity. Primarily, the focus of these studies has been on isolating and identifying the impact of *instructional expenditures*, which are assumed to most directly influence student outcomes such as standardized test scores. However, while the results of these analyses provide some limited support for the idea that “money matters” in the production of education, they have fallen short of forming a consensus that would override the uncertainty of previous studies.

In one such study, Sebold and Dato (1981) disaggregated expenditures into four categories, including (1) general education, (2) support services, (3) auxiliary programming, and (4) special education. They found some support for the hypothesis that general education expenditures were positively related to student outcomes, while the other three expenditure categories were not. Lopus (1990) also found support for the relationship between instructional expenditures and student outcomes, but her analysis was limited to high-school economics classes, and her most compelling results used proxy measures of instructional expenditures, such as teacher experience, class-size, and the quality of instructional materials. In contrast, Okpala, Okpala, and Smith (2001) found no link between instructional expenditures and student outcomes, though their analysis was limited to one rural county in North Carolina.

In a more rigorous study, Wenglinsky (1997) constructed a structural-equation model to examine the impact of various inputs on student outcomes. From his analysis he concluded that “... some spending measures play a role in student achievement while others do not” (p. 229). Specifically, Wenglinsky (1997) found both instructional expenditures and central administration expenditures to be positively related to student outcomes, while school administration and capital outlays were nonsignificant predictors. However, the link between instructional expenditures and student outcomes was mediated by class-size, making the claim of a direct relationship difficult to

support. Dee (2005) also found a significant relationship between instructional expenditures and student outcomes, but the same relationship prevailed for non-instructional expenditures as well, and the only measure of student outcomes considered was educational attainment/dropout.

Hypotheses

While the lack of consensus among these previous studies likely arises in part from their different sampling and modeling choices, it is also possible that the impact of instructional expenditures may be understated by a failure to consider the distinct goals of various instructional expenditure categories, such as the different impacts that regular program expenditures, special education expenditures, and vocational instruction expenditures have on student outcomes. By failing to account for these distinctions, researchers are effectively acting on the assumption that all instructional dollars have the same impact on educational productivity, which is unlikely to be true given the variety of educational functions represented in the broader “instructional expenditures” category.

For example, it would be reasonable to expect that regular program expenditures would impact standardized test scores (and the learning outcomes associated with them) more directly than those expenditures associated with special needs or vocational education. In contrast, while special education expenditures may help to improve standardized test scores for those special needs students with IEP’s who still participate in mainstream standardized testing, it might also be reasonably expected that many special education expenditures are directed at students who participate in alternative assessments. In this case the most direct impact of special education expenditures on student outcomes might be found in the area of alternative test scores. Likewise, vocational education expenditures may be expected to most directly influence other educational outcomes, such as dropout rates and career placement. By failing to properly disaggregate these expenditures or account for multiple outcome measures, previous studies may have inadvertently mistaken many of these expenditures for inefficiencies in the production function, when in fact they are simply advancing different educational goals.

Using detailed data from the Pennsylvania Department of Education, this article builds on the research discussed above by more thoroughly disaggregating instructional expenditures in an effort to better demonstrate the relationship between financial inputs and student outcomes. Specifically, this study considers the measured impact of instructional expenditures as a whole and then disaggregates the instructional expenditure category into (1) regular program expenditures, (2) special education instruction, (3) and vocational education instruction. It is hypothesized that this disaggregation of instructional expenditures will reveal a more robust relationship between regular program expenditures and standardized test-scores.

If this hypothesis holds true, then we would expect the parameter estimates for *regular program expenditures* to be larger than the parameter estimates for the aggregate *instructional expenditures* variable, which would be due to the removal of the weaker, non-existent, or countervailing relationships associated with expenditures in the other instructional subcategories. This would suggest that the impact of instructional expenditures is understated when the subcategories are not properly disaggregated.

Data and Methods

This study uses five consecutive years of district-level data for the Commonwealth of Pennsylvania’s public school districts. These data, which span from the 2006-07 school year through

the 2010-11 school year, have been collected from the Pennsylvania Department of Education's *Annual Financial Reports* and *PSSA Results*, as well as from the National Center for Education Statistics' *Common Core of Data*¹.

Pennsylvania's public school districts provide an ideal sample for production function analysis given the sharp disparities in per pupil expenditures that often exist across districts. Public school financing in Pennsylvania is more heavily dependent on local tax revenues than in many other states (PSFP, n.d.). Municipalities have extensive taxing authority and provide the majority of education funding, with state funding accounting for approximately a third of public education funds, and federal funding accounting for less than 10% (PSFP, n.d.). Verstegen (2011) classifies Pennsylvania's school funding system as a "foundation program", noting that in such cases, the prevalence of local taxing authority often results in significant expenditure disparities between poor and wealthier municipalities. It should be noted that the time-series analyzed in this study precede the passage of Pennsylvania's Opportunity Tax Credit Scholarship program, which was adopted in 2012; therefore the data are not influenced by that policy change. The data set is balanced panel of 349 school districts. We excluded the remaining 151 school districts to avoid confounding factors associated with missing data². The inclusion of these districts does not substantively change the results presented below.

While district level aggregation has previously been employed in a number of production functions models (e.g., Ferguson & Ladd, 1996; Gyimah-Brempong & Gyapong, 1991; Sebold & Dato, 1981), some objections have been raised to this approach on grounds of omitted variable biases (Hanushek, Rivkin, & Taylor, 1996). However, district-level analyses can be useful despite these methodological criticisms (Ferguson & Ladd, 1996), particularly for highlighting the general relationships between resource inputs and student outcomes. With that said, due to unobservable within-district variation, it is important to avoid any student-level interpretations of the parameter estimates, which would constitute a fallacy of division. Instead, particular attention should be paid to the larger, district-level patterns which may help to inform both policymakers and future researchers regarding the broader relationship between resource inputs and student outcomes.

Dependent Variables

For the purposes of this study, educational productivity is measured as the total percentage of students in each district whose scores are classified as either "Proficient" or "Advanced" on the Pennsylvania System of School Assessment (PSSA) exams. The PSSA exams are Pennsylvania's annual standardized tests, which (among other ends) are used in compliance with state and federal laws such as the No Child Left Behind Act of 2001. While standardized tests have been criticized as incomplete embodiments of educational goals (Barrow & Rouse, 2007), they remain the most commonly employed measure of student outcomes in production function research (Rice and Schwartz, 2008), due in part to their ease of availability in relation to other outcome variables. Some researchers have also suggested that test scores are a good proxy for future labor market returns (i.e. Murnane, Willett, Duhaldeborde, & Tyler, 2000), lending further credence, particularly from a human capital perspective, to their use as a measure of educational production.

¹ The five years of data analyzed in this study were chosen based on data availability limitations. Changes in the public reporting of PSSA results by the Pennsylvania Department of Education prohibited the use of a longer time-series without compromising measurement validity.

² Missing data results from data that are not reported because they fail to meet NCES data quality standards (see <https://nces.ed.gov/ccd/elsi/default.aspx?agree=0>). In this case, the excluded cases were primarily associated with missing data related to the race and SES variables, which may be withheld by some districts due to confidentiality concerns.

Table 1

Descriptive Statistics, Dependent Variables

Variable	<i>N</i>	\bar{x}	σ	min	max
Mathematics (% Advanced + Proficient)					
2006-07	349	72.79	9.82	20.4	92.3
2007-08	349	74.93	9.42	24.1	93.4
2008-09	349	76.55	8.62	32.1	94.4
2009-10	349	79.29	8.49	37.0	96.3
2010-11	349	79.89	8.38	39.0	96.6
Reading (% Advanced + Proficient)					
2006-07	349	71.33	10.17	26.7	91.9
2007-08	349	73.21	9.86	25.7	92.8
2008-09	349	74.35	9.37	34.2	93.1
2009-10	349	74.83	9.30	34.7	94.2
2010-11	349	76.32	9.63	35.5	94.9

Source: Pennsylvania Department of Education, PSSA Results.

This analysis considers exams in both Reading and Mathematics (analyzed separately), which are administered in grades 3 through 8, as well as grade 11, on an annual basis. Based on state-mandated performance levels, students are ranked in each subject as either (1) Advanced, (2) Proficient, (3) Basic, or (4) Below Basic. The performance levels are based on predetermined standards, not post-hoc percentiles. As the data in Table 1 show, PSSA outcomes improved on average over the five-year period in question, with the percentage of students classified as Advanced or Proficient rising by 9.75% in Mathematics and 6.99% in Reading. By using a district-wide percentage, this analysis attempts to mitigate some of the challenges associated with changing student cohorts in longitudinal studies, though the elimination of this problem is by no means complete. Furthermore, a district-wide measure of productivity is necessary since the available expenditure and control variables used in this analysis are also measured at the district level.

Independent Variables

Expenditure data were gathered from the Pennsylvania Department of Education's *Annual Financial Reports* for the 2006-07 through 2010-11 school years. These variables are analyzed at two different levels of aggregation. The first level disaggregates *Total Current Expenditures* into three exhaustive categories: (1) Instructional Expenditures, (2) Support Service Expenditures, and (3) Non-Instructional Expenditures. This disaggregation is consistent with that used in several of the studies discussed above (i.e. Dee, 2005; Sebold & Dato, 1981; Wenglinsky, 1997). The second level of aggregation builds on this approach by further disaggregating *Instructional Expenditures* into major subcategories: (1) Regular Instructional Programs, (2) Special Education Instruction, and (3) Vocational Education Instruction (while also retaining the Support Service and Non-Instructional categories). For ease of interpretation, all expenditure variables are measured in hundreds of dollars per pupil, and each expenditure variable is adjusted for inflation (\$2010).

As the descriptive statistics in Table 2 show, instructional expenditures increased consistently over the five-year period in question, with the exception of vocational expenditures, which on average remained relatively flat over that time. This makes these data an appropriate sample for this research, as detecting effects from time-variant predictors in a repeated measures model requires meaningful variation in the covariates over time. While the available data are aggregated at the district level, it should be noted that this approach limits our ability to observe variation in expenditure patterns across schools within the same district. However, the fact that Pennsylvania's school districts are organized and governed at the municipal-level makes this less of a concern than it would be in states employing a county-level aggregation, where socioeconomic conditions would be less homogenous within districts. It should also be noted that an analysis of collinearity diagnostics revealed no significant multicollinearity concerns with regard to the expenditure variables.

This study also accounts for organizational and student body control variables, which were obtained from the National Center for Education Statistics' *Common Core of Data*. The organizational control variables include Average Daily Membership (ADM) and teaching experience. ADM is employed as a proxy for district size. While economies of scale were hypothesized with regard to ADM in a previous study of Pennsylvania school districts (Klick, 2000), other researchers have actually found a negative relationship between student outcomes and district size, suggesting that smaller districts may be more efficient mechanisms of educational production (Fowler & Walberg, 1991; Robertson, 2007). The logged transformation of ADM is used in this analysis to account for the potential of diminishing economies of scale. Teaching experience is measured by "total years of service" as a district-level average for classroom teachers. Rivkin, Hanushek, & Kain (2005) note the critical role that quality teachers play in educational production, but operationalizing this construct has proven elusive. Teacher experience serves as an imperfect proxy for quality by accounting for factors such as professional experience and institutional knowledge.

The student-body control variables include (1) the percentage of students classified as "low-income", which is based on eligibility for the Federal Free and Reduced Lunch (FRL) program, (2) the percentage of students classified as "non-white", (3) the percentage of students classified as "Limited English Proficient" (LEP), and (4) the percentage of students on "Individualized Educational Plans" (IEP's). The inclusion of these control variables is important for several reasons. First, since the initial findings of the Coleman Report (1966), production function models have continued to identify a negative relationship between socioeconomic factors (such as poverty) and student achievement (Hanushek, 1997, 2003). In a previous study of Pennsylvania's school districts, Klick (2000) found poverty to be the most consistent predictor of student outcomes. Furthermore, it has also been shown that the costs associated with educating students from disadvantaged backgrounds, as well as students with special needs and limited English proficiency are significantly higher (Levin, 1989; Yinger 2001), which may influence the amount of money spent across districts as well as the measured student outcomes. It should be noted that the "non-white" and "low-income" variables were highly collinear, but both were retained in the analysis due to their theoretical significance and the fact that as control variables they are not directly related to the study's core hypotheses (Allison, 2012).

Statistical Models

After removing cases with incomplete data, the analysis was run on a sample of $n=349$ school districts (approximately 70% of the Commonwealth's 500 school districts). A fixed effects model of the following form was estimated to test the hypotheses outlined above:

$$y_{itk} = \alpha_i + \beta \text{Expend}_{mit} + \gamma Z_{it} + \theta T_t + u_{it} \quad \text{Eq. (2)}$$

where y is outcome k measured for district i in time period t ; α_i are district-specific intercepts; Expend is a vector of covariates measuring district expenditures across m instructional categories; Z is a vector of socio-economic, organizational, and environmental controls measured for district i in time period t ; and T is a vector of time dummies that capture common shocks to test scores across districts over time.

The fixed effects model specified in Equation 2 has several advantages over cross-sectional or pooled regression models. Foremost among them is that this strategy removes the bias in our estimates of β from fixed confounders omitted from the model. This is accomplished by discarding the between-district variation in the outcome and explanatory variables “contaminated” by unobserved fixed factors that may explain both district-level expenditures and test scores (Allison 2009). This is an attractive feature for this analysis as the heterogeneity in district expenditure levels within each of the m categories is likely correlated with several relevant unobserved factors that are plausibly fixed over the five-year observation period, including the local tax base of districts in the sample. In effect, the estimates in Equation 2 are derived from the variation in these measures over time within the district. The descriptive statistics reported in Table 2 confirm the existence of variation in these measures within districts over time, which is necessary to efficiently estimate the coefficients in Equation 2.

Table 2
Descriptive Statistics, Independent Variables

	<u>Academic Years</u>									
	<u>2006-07</u>		<u>2007-08</u>		<u>2008-09</u>		<u>2009-10</u>		<u>2010-11</u>	
	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
Instructional Expenditures [†]	72.38	11.61	74.59	12.37	75.39	12.42	77.69	12.59	79.38	13.00
Regular Programs [†]	50.40	8.04	52.78	8.46	53.03	8.32	54.17	8.44	55.46	8.82
Special Education [†]	15.06	4.88	15.67	5.09	16.23	5.20	17.23	5.28	17.77	5.34
Vocational Instruction [†]	3.93	2.08	3.91	2.05	3.94	2.05	4.01	2.04	3.96	2.06
Support Service Expenditures [†]	37.62	7.66	38.82	7.80	39.67	7.77	39.64	7.49	40.28	7.73
Non-Instructional Expenditures [†]	2.21	1.08	2.27	1.14	2.27	1.11	2.30	1.13	2.90	0.97
Average Daily Membership (ADM)	4346.30	11087.84	4337.39	11285.54	4304.87	11203.89	4291.82	11277.99	4273.82	11152.51
Percent Free/Reduced Lunch Students	27.38	16.05	28.83	16.72	30.56	16.89	32.89	17.54	33.79	17.65
Percent Non-White Students	14.21	17.16	13.94	17.62	14.46	17.93	15.17	18.17	15.92	18.38
Teaching Experience in Years	13.93	2.09	13.93	2.09	13.61	2.00	13.45	1.84	13.52	1.91
Percent LEP Students	1.26	2.10	1.37	2.28	1.40	2.42	1.36	2.39	1.36	2.31
Percent IEP Students	16.20	3.15	16.32	3.14	16.48	3.24	16.44	3.30	16.39	3.30

Source: Pennsylvania Department of Education, Annual Financial Reports.

[†] All expenditure variables measured in hundreds of dollars per pupil and adjusted for inflation (shown as 2010 dollars).

Results

Table 3 contains results for the production function models, with separate models run for standardized Mathematics and Reading tests. In each case, the models were initially run with *Total Current Expenditures* disaggregated into three exhaustive categories, including (1) Instructional Expenditures, (2) Support Service Expenditures, and (3) Non-Instructional Expenditures. Then the models were rerun with the Instructional Expenditures category disaggregated into (1) Regular Program Expenditures, (2) Special Education Expenditures, and (3) Vocational Education Expenditures.

Overall, the results show moderate support for the study's primary hypothesis that the impact of instructional expenditures on student outcomes is understated when instructional subcategories are not properly considered. The results support this hypothesis most directly in the case of standardized Mathematics tests, though less so in the case of standardized Reading tests. For example, in Model 1, the aggregate measure of *instructional expenditures* is positively related to standardized Math scores, with a parameter estimate of 0.143 ($p \leq .01$). However, once instructional expenditures are disaggregated in the second model, the parameter estimate for *regular program expenditures* is 0.151 ($p \leq .01$), roughly 6% larger than the overall parameter estimate for *instructional expenditures* in the aggregate model. The parameter estimates for both *special education expenditures* and *vocational instruction expenditures* are statistically nonsignificant in this instance. Thus, the relationship between regular program expenditures and test scores appears to be slightly attenuated when this amount is combined with the other instructional expenditure categories, which appear to have a less direct impact on standardized Math scores.

Table 3
Repeated Measures Models, %Advanced + Proficient (N= 1745)

	<u>Mathematics</u>		<u>Reading</u>	
	I	II	III	IV
Instructional Expenditures†	.143***	-	.091**	-
Regular Programs†	-	.151***	-	.064**
Special Education†	-	.070	-	.060
Vocational Education†	-	.135	-	.269**
Support Service Expenditures	.008	.015	.028	.036
Non-Instructional Expenditures	-.155	-.162	.022	-.001
ADM (log)	-5.137	-5.446	2.558	1.667
Free/Reduced Lunch Eligible (%)	.067**	.066**	.035	.036
Non-White (%)	-.128	-.132	-.164**	-.166**
Teachers Average Years of Experience	.075	.080	.108	.123
% LEP Students	-.154	-.180	-.235*	-.240*
% IEP Students	-.151	-.147	-.042	-.043
Year Fixed Effects				
2007	1.690***	1.600***	1.597***	1.599***
2008	3.139***	3.078***	2.738***	2.744***
2009	5.467***	5.488***	3.067***	3.128***
2010	5.904***	5.942***	4.471***	4.585***
Constant	105.304**	108.690**	43.963	52.080*

* $p \leq .10$; ** $p \leq .05$, *** $p \leq .001$

† All expenditure variables are measured in hundreds of dollars per pupil.

In the case of standardized Reading tests, the aggregate measure of *instructional expenditures* was again statistically significant and positively related to improved student outcomes (Model 3), with a parameter estimate of 0.091 ($p \leq .05$). Once the *instructional expenditures* variable was disaggregated into subcategories in Model 4, the overall explanatory power of the expenditure variables did increase, however the parameter estimate for *regular program expenditures* was slightly smaller ($\beta = 0.064$; $p \leq .05$) than the original estimate for *instructional expenditures*. Surprisingly, the largest parameter estimate was associated with *vocational instruction expenditures*, where the β coefficient was 0.269 ($p \leq .05$), nearly three times larger than the original parameter estimate for instructional expenditures. As in the case of Mathematics, the parameter estimates for *special education expenditures* remained statistically nonsignificant. These findings raise two important considerations, including the size of the relationship between vocational instruction expenditures and standardized Reading scores, as well as the differential impact of regular program expenditures on Mathematics and Reading scores.

While the magnitude of the relationship between vocational expenditures and standardized Reading tests is unexpectedly large, this finding is not entirely surprising. As several scholars have

noted, reading comprehension is essential to success in the workplace. For example, Garton (2012) argues that

... illiteracy is simply not an option for CTE (career and technical education) students. Today's workplace requires the ability to read and absorb technical manuals, understand and program computers, write and respond to memos on technical and professional matters, and interpret tables of instructions. In fact, CTE texts can contain very difficult content, on par with or more difficult than traditional academic courses. (p.2)

Furthermore, vocational education may create an environment that is more conducive to the cultivation of literacy skills insofar as students are more engaged in those materials which directly apply to their areas of personal and professional interest (Garton, 2012). Based on this reasoning, a variety of efforts have been made to integrate reading comprehension and literacy skills into the vocational education curriculum (i.e. Darvin, 2005; Schneider & Foot, 2013). Add to this the fact that many vocational education students receive a "double-dose" of reading instruction, both in the traditional classroom and in their more technical vocational curriculum, and this finding is not altogether unexpected. However, while these arguments may help to explain the positive relationship between vocational instruction expenditures and standardized Reading scores, the size of the observed relationship is still surprising and may suggest some unique features of the PSSA exams which favor or reward technical reading comprehension skills. Either way, further investigation of this relationship is warranted in order to determine if this finding is externally reliable and generalizable across various school and policy environments.

With regard to the different observed impact of *regular program expenditures* in the case of Mathematics and Reading, these results may be due in part to the diverse educational contexts in which these subjects are taught. For example, documented teacher shortages in the STEM disciplines (i.e. Hutchinson, 2012; U.S. Dept. of Education, 2002) may mean that the costs for recruiting and retaining quality Math teachers are higher than in the case of Reading. Furthermore, the increasing prominence of classroom technology in Mathematics instruction (i.e. NCTM 2011) may also contribute to the more pronounced relationship between educational expenditures and student proficiency in Mathematics. If this is the case, then the optimal specification of production function models may vary across subject matters and testing areas.

Despite these considerations, the overall results do suggest that disaggregated approaches to measuring instructional expenditures enhances the explanatory power of production function models and at least moderately improves our understanding of the relationship between public expenditures and educational productivity. While the parameter estimates associated with expenditure variables tend to be small in these models, this is of less overall importance than how the disaggregation of instructional expenditures influences their measured impact on student outcomes. Historically, effect sizes have varied across production function estimates, and this may be largely due to differences in sampling and the units of analysis considered (i.e. districts, schools, and students). In this case, the larger measured effect of regular instructional programming for Mathematics, and even the greater explanatory power of the disaggregated Reading model, suggests that previous production functions based on aggregate measures of expenditures (whether total or instructional) may have understated the impact of educational expenditures on student outcomes. It should also be pointed out that the parameter estimates for support service and non-instructional expenditures were not statistically significant predictors of student outcomes in any of the models, which is consistent with the findings of some prior research (i.e. Sebold & Dato 1981).

While the additional control variables were not directly relevant to this study's core hypotheses, it should be noted that they were largely non-significant predictors of student outcomes in these models. The percentage of non-white students is statistically significant and negatively related to educational outcomes in the two models evaluating reading scores, regardless of the level of financial aggregation. Parameter estimates for the low-income variable were positive in each case, which runs contrary to the findings of previous studies, but this is most likely the result of high multicollinearity between it and the percent non-white variable, as mentioned above. While teacher experience was positively related to student outcomes, these increases were not statistically significant in any of the models. Districts with higher percentages of LEP students did tend to have lower reading scores but this relationship was only statistically significant at the 0.10 level. The percentage of IEP students was not significantly associated with changes in either Math or Reading scores. This may again be due to high multicollinearity among the control variables.

Discussion

Several implications for both policy and research arise from the results of this analysis. First and foremost, these results suggest that the impact of instructional expenditures may have been understated in many prior production function studies. This argument is supported most directly by the parameter estimates associated with *regular program expenditures* for standardized Mathematics tests. Once *instructional expenditures* are disaggregated into more specific subcategories, the parameter estimates for *regular programs* are 6% larger than the estimates for the aggregate *instructional expenditures* category. This finding suggests that regular programming expenditures may be a more appropriate measure of what instructional expenditures were intended to capture in previous studies, namely the relationship between instructional investments and traditionally measured student outcomes (i.e. standardized test scores). This dynamic is likely due to the fact that instructional dollars associated with *special education* and *vocational instruction* do not equally influence standardized Math scores, which may attenuate the observed relationship between an aggregated expenditure measure and this outcome.

For production function researchers, these findings not only suggest a need for greater specification with regard to expenditure variables, but also with regard to the measurement of student outcomes. Intuitively, we might have expected that dollars spent on special education would not influence standardized test scores in the same manner as those spent directly on regular educational programming. However, negative coefficients and non-significance do not mean that these expenditures are therefore inefficient or “unproductive”; they are simply focused toward different educational outcomes. In order to properly gauge the effectiveness of these expenditures, they need to be examined in relation to the appropriate measures of special needs and vocational student outcomes. In one sense, the disaggregation of instructional expenditures along these lines is simply an acknowledgement of the fact that different instructional investments serve distinct educational goals. Through this lens, the focus of future analyses may be more accurately seen, not as determining whether inputs influence production, but rather understanding how different inputs influence a variety of outcomes in different ways. The observed relationship between vocational instruction expenditures and standardized Reading scores would be one important area of consideration. The policy goals that arise from such an understanding would likely be more nuanced, though in the long run they would presumably be more effective.

Finally, from a policy perspective, these results raise some legitimate objections to the “money doesn't matter” mantra which has become popular in some policy circles (for discussion and examples see Baker, 2012). To the extent that this “non-relationship” may be heralded on the

basis of an incomplete production function, it is premature at best. Policies based on the assumption of a non-relationship between educational expenditures and student outcomes could have negative and significant consequences for current and future students. At this point, additional analysis should be conducted in order to more accurately understand the relationship between specific instructional investments and their appropriate outcome measures. Ideally, this should include analyses across a variety of state policy settings, with consideration given to varying units of analysis (i.e. districts, schools, and students). A lack of consensus around how best to measure educational outcomes, as well as challenges with data availability, make this easier said than done in many instances, but this further analysis is essential to well informed education finance policies.

Conclusion

Despite the conclusions of the Coleman Report (1966) and many subsequent studies (i.e. Hanushek, 1997, 2003), education economists and policy analysts have continued to pursue more robust specifications of the production function model, holding to the premise that resource inputs *should* positively impact educational productivity. While the findings of this analysis do not definitively affirm this premise, they do suggest that these basic intuitions have merit and that the goals of production function research may continue to be advanced through the pursuit of more granular data and improved model specifications. In this pursuit, greater attention should be paid to linking specific instructional investments with the appropriate educational outcomes, which may hold promise for improving our understanding of the relationship between public expenditures and educational productivity.

References

- Allison, P. D. (2009). *Fixed Effects Regression Models (Quantitative Applications in the Social Sciences)*. Sage Publications.
- Allison, P. D. (2012, September 10). When can you safely ignore multicollinearity? Retrieved from <http://statisticalhorizons.com/multicollinearity>
- Baird, K. (2011). Assessing why some students learn math in high school: How useful are student-level test results? *Educational Policy* 25, 784-809.
- Baker, B. B. (2012). Revisiting the age old question: Does money matter in education? *The Albert Shanker Institute*. Available at http://www.shankerinstitute.org/images/doesmoneymatter_final.pdf
- Barrow, L., & Rouse, C. E. (2007). Causality, causality, causality: The View of education inputs and outputs from economics. In S. H. Fuhrman, D. K. Cohen, & F. Mosher (Eds.) *The State of Education Policy Research*. Mahwah, NJ: Lawrence Erlbaum.
- Bowles, S., & Levin, H. M. (1968). The Determinants of scholastic achievement: An Appraisal of some recent evidence. *The Journal of Human Resources*, 3, 3-24.
- Coleman, J. S. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.
- Darvin, J. (2006). On reading recipes and racing forms: The Literacy practices and perceptions of vocational educators. *Journal of Adolescent and Adult Literacy*, 50, 10-18.
- Dee, T. S. (2005). Expense preference and student achievement in school districts. *Eastern Economic Journal*, 31, 23-44.
- Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An Analysis of

- Alabama schools. In Helen F. Ladd (Ed.). *Holding Schools Accountable* (pp. 265-298). Washington, D.C.: The Brookings Institute.
- Fowler, W. J. & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis* 13, 189-202.
- Garton, R. (2012, August 13). Literacy for life: The Role of career and technical education in reading proficiency. Retrieved from <http://www.shankerinstitute.org/blog/literacy-life-role-career-and-technical-education-reading-proficiency>
- Greenwald, R., Hedges, L., & Laine, R. D. (1996). The Effect of school resources on student achievement. *Review of Educational Research*, 66, 361-396.
- Gyimah-Brempong, K., & Gyapong, A. O. (1991). Characteristics of education production functions: An application of canonical regression analysis. *Economics of Education Review* 10, 7-17.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An Update. *Educational Evaluation and Policy Analysis*, 19, 141-164.
- Hanushek, E. A. (2003). The Failure of input based schooling policies. *The Economic Journal* 113, 64-98.
- Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics* 78, 611-627.
- Hutchinson, L. F. (2012). Addressing the STEM teacher shortage in American schools: Ways to recruit and retain effective STEM teachers. *Action in Teacher Education*, 34, 541-550.
- Klick, J. (2000). Do dollars make a difference? The Relationship between expenditures and test scores in Pennsylvania's public schools. *American Economist*, 44, 81-87.
- Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal*, 113, F34-F63.
- Levin, H. M. (1989). Financing the education of at risk students. *Educational Evaluation and Policy Analysis*, 11, 47-60.
- Lopus, J. S. (1990). Do additional expenditures increase achievement in the high school economics class. *Journal of Economic Education*, 21, 277-286.
- Marion, R., & Flanigan, J. (2001). Evolution and punctuation of theories of educational expenditures and student outcomes. *Journal of Education Finance*, 26, 239-258.
- Monk, D. H. (1989). The Education production function: Its Evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, 11, 31-45.
- Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19, 547-568.
- National Council of Teachers of Mathematics (NCTM). (2011). *Technology in teaching and learning mathematics*. Reston, VA: NCTM.
- O'Connell Smith, M. (2004). How the apportionment of district money is associated with differences in the grade 8 mathematics and reading achievement scores for Minnesota students. *Journal of Education Finance* 29, 299-314.
- Okpala, C. O., Okpala, A. O., & Smith, F. E. (1995). Parental involvement, instructional expenditures, family socioeconomic attributes, and student achievement. *The Journal of Educational Research*, 95, 110-115.
- Pennsylvania School Funding Project (PSFP). (n.d.). *Pennsylvania's Historic School Funding Problem*. Harrisburg, PA: PSFP. <http://www.paschoolfunding.org/the-problem/pennsylvania%E2%80%99s-historic-school-funding-problem/>
- Rice, J. K., & Schwartz, A. E. Toward an understanding of productivity in education. In

- H. F. Ladd & E. B. Fiske (Eds.) *Handbook of Research in Education Finance and Policy* (pp. 131-145). New York: Routledge.
- Rivkin, S. G., Hanushek, E.A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417-458.
- Robertson, F. W. (2007). Economies of scale for large school districts: A National study with local implications. *The Social Science Journal*, 44, 620-629.
- Sebold, F. D., & Dato, W. (1981). School funding and student achievement: An Empirical analysis. *Public Finance Review*, 9, 91-105.
- Schneider, J. L., & Foot, R. (2013). Teaching strategies to support vocational education students' reading literacy. *The Clearing House*, 86, 32-36.
- Summers, A. A., & Wolfe, B. L. (1977). Do schools make a difference? *The American Economic Review*, 67, 639-652.
- U.S. Department of Education. (2002). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality*. Washington, DC: U.S. Department of Education, Office of Postsecondary Education.
- Verstegen, D. A. (2011). Public education finance systems in the united states and funding policies for populations with special educational needs. *Education Policy Analysis Archives* 19, <http://dx.doi.org/10.14507/epaa.v19n21.2011>
- Wenglinsky, H. (1997). How money matters: The Effect of school district spending on academic achievement. *Sociology of Education*, 70, 221-237.
- Yinger, J. (2001). Fixing New York State's education aid dinosaur: A proposal. Center for Policy Research, Paper 22. <http://surface.syr.edu/cpr/22>

About the Authors

Stephen R. Neely

University of South Florida

srneely@usf.edu

Dr. Neely holds a Ph.D. in Public Administration from North Carolina State University. He is an assistant professor at the University of South Florida. He currently teaches courses in research methods and public policy at the University of South Florida's School of Public Affairs. He conducts research in the areas of public affairs education and K-12 education policy.

Jeffrey Diebold

North Carolina State University

jcdiebol@ncsu.edu

Dr. Diebold is an assistant professor at North Carolina State University where he teaches public policy analysis and quantitative research methods. His research interests include social welfare policy, pension and retirement policy, and research design.